THE ACTIN I INTRON—A PHYLOGENETICALLY INFORMATIVE DNA REGION IN *CLEMATIS* (RANUNCULACEAE)

Jonathan M. Slomba, James R. Garey, and Frederick B. Essig

Department of Biology
University of South Florida
Tampa, Florida 33620, U.S.A.
essig@chuma1.cas.usf.edu

ABSTRACT

As part of a search for DNA regions suitable for phylogenetic analysis in the genus *Clematis* (Ranunculaceae), the nuclear-encoded Actin I intron was employed in a preliminary sampling of a small number of species selected to represent the major subdivisions of the genus. This DNA region was found to be more informative and to provide a more robust phylogenetic tree than chloroplast DNA regions. Trees generated were consistent with phylogenetic hypotheses based on morphology and with published molecular analyses. Moreover, three species of the section Crispae (subgenus Viorna) native to Florida, considered closely related on morphological grounds, formed a well-supported clade, and were distinguishable from one another. This suggests that this DNA region might be a useful tool for distinguishing groups of closely-related species, individual species, and possibly hybrids.

RESUMEN

Como parte de una búsqueda de regiones de DNA apropiadas para análisis filogenéticos en el género *Clematis* (Ranunculaceae), se empleó el intrón Actin I nuclear en un muestreo preliminar de un pequeño número de especies seleccionadas para representar las grandes subdivisiones del género. Esta región de DNA se encontró que era más informativa y daba un árbol filogenético más consistente que las regiones de DNA plastidial. Los árboles generados eran consistentes con la hipótesis filogenética basada en la morfología y con los análisis moleculares publicados. Además, tres especies de la sección Crispae (subgénero Viorna) nativas de Florida, consideradas fuertemente emparentadas en aspectos morfológicos, formaron un clado muy coherente, y eran distinguibles una de otra. Esto sugiere que esta región del DNA puede ser un instrumento útil para diferenciar grupos de especies muy emparentadas, especies individuales, y posiblemente híbridos.

INTRODUCTION

The infrageneric classification of *Clematis* (Ranunculaceae), a genus of more than 300 species distributed worldwide, has been uncertain pending definitive phylogenetic studies. Traditional classifications have relied primarily on floral characters for the major divisions of the genus (as in Tamura 1967). However, characters of seedling and juvenile morphology have been cited in recent decades as supporting a fundamental division in the infrageneric classification of the genus (Tamura 1987; Essig 1991; Miikeda et al. 1999). A specialized syndrome of seedling and vegetative characters, featuring a suppressed hypocotyl

SIDA 21(2): 879–886. 2004

and opposite seedling leaves ("Type II," Essig 1991, "opposite" in Miikeda et al. 1999, see also Appendix 1) appears to have arisen from the more general Ranunculaceous type featuring an elongate hypocotyl and alternate seedling leaves ("Type I," or "alternate"), but it has been uncertain whether this morphological complex has arisen just once or more than once, as it is found in species formerly placed in different subgenera.

We have been seeking appropriate molecular tools with which to resolve these and other phylogenetic questions within *Clematis*. The use of DNA sequencing techniques has thus far been of limited success. Miikeda et al. (1999) utilized several chloroplast genes, including matK (maturase-encoding gene), trnK (UUU) intron, trnL (UAA) intron, the intergenic spacer between trnL and trnF (GAA), and the intergenic spacer between rbcL and atpB. Employing approximately 4,400 bp. of sequence for the eight taxa included in the study, the team produced a tree that was consistent with Essig's proposal, but was weakly resolved. Our own efforts with chloroplast DNA also produced weak results.

A search for alternative tools led us to consider some nuclear DNA regions, which are expected to be more informative than chloroplast non-coding regions in determining species level phylogenies because the nuclear genome has a substitution rate 5 to 10 times faster than the chloroplast genome (Li 1997). In particular, we have focused on a non-coding intron of the Actin I gene. Actin is one of the components of the cellular cytoskeleton, and is produced through the activity of a large multigene family (Moniz de Sa & Drouin 1996). The location of the Actin I intron is highly conserved across all angiosperm families, making the development of primers suitable for the PCR amplification of the intron possible. A preliminary test of this DNA region as a phylogenetic tool in *Clematis* was conducted using a small sampling of species representing the major subdivisions of the genus.

MATERIALS AND METHODS

Materials were obtained from the Chicago Botanical Garden and the University of South Florida Botanical Garden (Table 1). Samples were selected to represent the major subdivisions of the genus *Clematis*. They include several species with Type I vegetative morphology, and several with Type II vegetative morphology, while also representing the traditional sections *Clematis* and *Viorna* (as in Tamura 1967), and the rearranged sections (elevated to subgenera) of Tamura (1987) (Table1). The traditional sections each contained subsections with Type I and Type II morphologies, and in his revision, Tamura (1987) reorganized his classification to reflect those different morphologies. Note that two of the species included in this study were realigned in that taxonomic shift. *Clematis terniflora* in subsection *Rectae* was formerly included in section Clematis, while *C. stans*, in subsection *Tubiflorae* was formerly included in section *Viorna*. This analysis is thus a preliminary test of that taxonomic revision.

TABLE 1. Taxa included in the analysis (with seedling morphology type indicated as I or II).

Species	Classification (Tamura 1987)	Voucher
Clematis reticulata Walt.	(Viorna: Crispae II)	*Arias 71* (USF)
Clematis crispa L.	(Viorna. Crispae II)	*Essig 011001-6* (USF)
Clematis baldwinii Torr. & A. Gray	(Viorna: Crispae II)	*Essig 011001-7* (USF)
Clematis terniflora DC	(Flammula: Rectae II)	*Essig 860904-1* (USF)
Clematis virginiana L.	(Clematis: Dioicae I)	Chicago B. G. acc. # 356-81
Clematis stans Sieb. & Zucc.	(Campanella: Tubulosae I)	*Essig 011001-3* (USF)
Anemone pulsatilla var. vulgaris L.	(outgroup I)	*Essig 020305-2* (USF)

The samples for this study also includes three species native to Florida that on morphological grounds appear to be closely related. Their inclusion provides a test of the resolving power of the Actin I DNA sequence.

A species of *Anemone* (*A. pulsatilla* var. *vulgaris* L) was chosen as the outgroup. *Anemone* has traditionally been identified as closely related to *Clematis*, and *A. pulsatilla* shares with *Clematis* the very distinctive elongate styles of the mature achenes. A number of recent phylogenetic studies (Johansson & Jensen 1993; Hoot 1995, and Kosuge et al. 1995) have identified a clade that includes *Clematis* along with *Anemone*, *Pulsatilla* (sometimes treated as a segregate of *Anemone*), *Knowltonia*, *Hepatica*, and sometimes *Ranunculus* and/or *Trautvettaria*. A more comprehensive study will include more of these genera as outgroups. An unnamed species of *Anemone* was also used as the outgroup in the study by Miikeda et al. (1999).

Angiosperm Actin (ac1) gene sequences from a broad range of taxa were obtained from Genbank (*Arabidopsis*, *Zea*, *Oryza*, and *Glycine* (accession #'s M20016, J01238, X15865, and J01298, respectively). These sequences were aligned using Clustal X (Thompson et al. 1994; Higgins et al. 1996). The primer sequences were selected from the alignment by anchoring the forward primer in a highly conserved (relatively guanine and cytosine rich [48%]) coding region just downstream of the intron. The reverse primer was anchored in a highly conserved coding region just upstream of the intron (Table 2). This primer set corresponds to a region of approximately 300 nucleotides in the taxa listed above.

Some specimens were deep frozen at -80° C before use, others were prepared immediately for extraction. Total genomic DNA was extracted from leaf samples following the modified CTAB protocol developed by Doyle & Doyle (1987).

Polymerase chain reaction (PCR) was carried out on all extracted DNA samples using primers for the gene regions shown in Table 2. PCR reactions (amplification) were carried out in 100 µL volumes, using a taq polymerase kit from Enzypol (Boulder CO), following their instructions. Thermal cycling parameters were the same for all species: 1 min. initial denature at 95°C, followed by 35

TABLE 2. PCR primers for Actin genes.

Actin I forward:	CCC	GAA	TTC	CTT	GTT	TGC	GAC	AAT	GGA	AC
Actin I reverse:	CCC	GAA	TTC	ACA	ATT	CCA	TGC	TCA	AT	

cycles of 15 sec. at 95°C, annealing at 48 °C for 30 sec., a 90 sec. extension at 72°C, followed by a 10 min. hold at 72°C, and then a final hold at 4°C.

The ~300 bp PCR product of the Actin I intron was gel purified and cloned into a pBluescript vector and transformed into DH5μ *Escherichia coli* cells (Gibco, Carlsbad CA). White colonies were picked from bacterial plates, grown in small cultures with ampicillin and sequencing template prepared using the alkaline lysis method (Ausubel 2000). After purification, cycle sequencing reactions were carried out in 200μL thin-walled capped tubes using a Perkin-Elmer (Foster City, CA) DYEnamic ET terminator cycle sequencing kit with an ABI Model 310 genetic analyzer.

Phylogenetic trees were constructed from DNA sequence alignments generated by Clustal X (Thompson et al. 1994; Higgins et al. 1996) and modified by hand using Genedoc (Nicholas et al. 1997). Neighbor-joining trees from molecular data were made using MEGA 2.1 software (Kumar et al. date) using Kimura 2-parameter distances with 1000 bootstrap replicates. For comparison, a brief morphological analysis incorporating vegetative and floral characters used in recent classifications was carried out (Fig. 1B, and Appendices 1 & 2). The morphological tree was recovered using PAUP 4.0b10 (Swofford 2001) using an heuristic search with default parameters and 200 bootstrap replicates.

RESULTS AND DISCUSSION

The Actin intron sequences yielded an alignment of 316 bp with 27 variable sites, 17 parsimony informative sites and 28 sites with gaps. The alignment was relatively unambiguous and resulted in a neighbor-joining tree (Fig. 1A) with two well-supported clades that coincide with Type I and Type II seedling morphology as described by Essig (1991). A maximum parsimony tree based on morphology (Fig. 1B) recovered the type II clade, and illustrates the plesiomorphic distribution of the Type I character syndrome. The molecular analysis produced a strongly supported (bootstrap value 100%) derived clade containing *Clematis crispa, C. reticulata* and *C. baldwinii*. These are morphologically similar species native to the southeastern U.S. belonging to the traditional group *Crispae* (variously designated as a section or subsection) in subgenus *Viorna*, and exhibiting type II morphology. There is a strong sister group relationship (97%) between this group and *C. terniflora*, a Eurasian species also with Type II morphology, but with panicles of small whitish flowers—a reproductive morphology syndrome it shares with Type I members of the traditional subgenus *Clematis*. *Clematis virginiana* and *C. stans* form a separate well-sup-

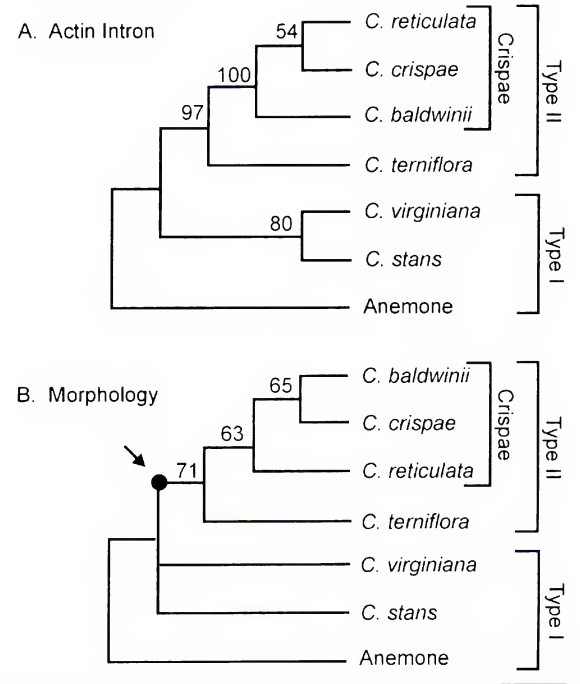

FIG. 1. **A.** Neighbor Joining tree from the Actin intron alignment. A maximum parsimony analysis (not shown) also found the Type II and *Crispae* nodes but with 90 and 99 percent bootstrap values respectively. **B.** Maximum parsimony tree from morphological data (see Appendices 1 and 2). Type I and Type II are syndromes of morphological characters, primarily of seedlings, as described in Essig (1991). The arrow on B indicates the origin of the Type II syndrome. *Crispae* is the sectional name for North American species of subgenus *Viorna* (Tamura 1987).

ported clade (80%) and both have Type I seedling morphology, despite having different floral morphologies.

The results of this preliminary analysis are consistent with those obtained by Miikeda et al. (1999). The species used in the two studies were different, but representative of the same infrageneric taxa. The results are also supportive of Tamura's (1987) revised classification, and Essig's (1991) proposal that taxa with type II seedling morphology represent a monophyletic clade and might be placed together in a major infrageneric division of the genus.

The results also confirm the close relationship of the 3 species of the subsection *Crispae* occurring in the southeastern U.S.A., and appear to resolve those species from one another. The relationship of the three species was slightly different in the DNA analysis from that in the morphological analysis, or from what one would expect through conventional taxonomic analysis. More extensive sampling within species is needed, along with analysis of additional DNA regions, to fully evaluate the resolving power of the Actin I intron region at this level.

Another discrepancy between the two analyses is the sister group relationship between *C. virginiana* and *C. stans* found in the molecular tree but not in the morphological tree. Too few taxa were included in this study to draw any conclusions about the deeper branches in the genus, however. A great many more taxa with both Type 1 and Type 2 morphologies exist. A more complete analysis will include a great many more of the species of this large genus, and in particular, as many of the recognized infrageneric taxa (sections, subsections) as possible, along with a comprehensive morphological analysis, in order to fully understand the phylogeny of this genus and develop a definitive infrageneric classification.

In conclusion, the results of this preliminary analysis are consistent with taxonomic concepts based on morphology and with other DNA-based analyses, and also appear to discriminate among fairly closely related species. Therefore, it appears that the Actin I gene region will be a very useful tool for the analysis of infrageneric relationships in *Clematis*, and likely in other angiosperm genera.

APPENDIX 1

Characters used in morphological analysis. Note: characters 1–4 are the primary features distinguishing the Type I (0) from the Type II (1) syndrome; characters 6–8 are the floral characters traditionally cited in distinguishing subgenus *Clematis* from subgenus *Viorna*.

1. Seedlings with hypocotyl elongate (0) vs hypocotyl suppressed (1)
2. Seedling leaves alternate (0) vs leaves opposite (1)
3. Eophylls 3-lobed (0) vs eophylls elliptic (1)
4. Leaves dentate (0) vs leaves entire (1)
5. Stems erect (0) vs stems vining (1)

6. Flowers with sepals spreading to reflexed from the base(0) vs flowers tubular, urceolate or campanulate with sepals spreading at the tips(1) vs flowers campanulate with strongly reflexed limbs (2)
7. Flowers colored (0) vs flowers white to cream (1)
8. Stamens with filaments glabrous (0) vs filaments hirsute (1)
9. Achenes narrow, turgid (0) vs achenes broad, flattened (1)

APPENDIX 2. SPECIES/MORPHOLOGICAL CHARACTER MATRIX

	1	2	3	4	5	6	7	8	9
Anemone pulsatilla	0	0	0	0	0	0	0	0	0
Clematis stans	0	0	0	0	1	1	0	1	0
Clematis virginiana	0	0	0	0	0	0	1	0	0
Clematis terniflora	1	1	1	1	0	0	1	0	1
Clematis reticulata	1	1	1	1	0	1	0	1	1
Clematis crispa	1	1	1	1	0	2	0	1	1
Clematis baldwinii	1	1	1	1	1	2	0	1	1

ACKNOWLEDGMENTS

This publication was derived from the Master's thesis of the first author. We acknowledge the support of NSF grants CHE-0221834 and DEB-0344372. We thank the Chicago Botanic Garden for providing materials. We thank Nancy Moreno and John Syring for providing valuable review comments.

REFERENCES

Ausubel, F.M., R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidmann, J.A. Smith, and K. Struhl. 2000. Current protocols in molecular biology. Greene Publishing Associates, NY.

Doyle, J.J. and J.E. Doyle. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf material. Phytochemistry J. 19:11–15.

Essig, F.B. 1991. Seedling morphology in *Clematis* (Ranunculaceae) and its taxonomic implications. Sida 14:377–390.

Higgins, D.G., J.D. Thompson, and T.J. Gibson. 1996. Using CLUSTAL for multiple sequence alignments. Meth. Enzymol. 266:383–402.

Hoot, S.B. 1991. Phylogeny of the Ranunculaceae based on epidermal microcharacters and macromorphology. Systematic Bot. 16:741–755.

Hoot, S.B. 1995. Phylgenetic relationships in *Anemone* (Ranunculaceae) based on DNA restriction site variation and morphology. Pl. Syst. Evol., suppl. 9:195–200.

Johansson, J.T. and R.K. Jensen. Chloroplast DNA variation and phylogeny inof the Ranunculaceae. Pl. Syst. Evol. 187:29–49.

Kosuge, K., K. Sawada, T. Denda, and K. Watanabe. 1995. Phylogenetic relationships of some genera in the Ranunculaceae based on alcohol dehydrogenase genes. Pl. Syst. Evol., Suppl. 9:263–271.

Kumar, S., K. Tamura, I.B. Jakobsen, and M. Nei. 2001. MEGA2: Molecular evolutionary genetics analysis software, Arizona State University, Tempe.

Li, W.H. 1997. Molecular evolution. Sinauer Assoc. Inc., Sunderland, MA.

Miikeda, O.S. Koga, T. Handa, and T. Yukawa. 1999. Subgeneric relationships in *Clematis* (Ranunculaceae) by DNA sequences. In: S. Andrews, A.C. Leslie, and C. Alexander, eds. Taxonomy of cultivated plants. Third International Symposium. Royal Botanic Gardens. Kew. Pp. 355–358.

Moniz de Sa, M. and G. Drouin. 1996 . Phylogeny and substitution rates of angiosperm actin genes. Molec. Bio. Evol. 13:1198–1212.

Nicholas, K.B., H.B. Nicholas, and D.W. Deerfield II. 1997. GeneDoc: analysis and visualization of genetic variation, Embnew. News 4:14.

Tamura, M. 1967. Morphology, ecology and phylogeny of the Ranunculaceae VII. Sci. Rep. Osaka Unv. 16–2:21–43.

Tamura, M. 1987. A classification of the genus *Clematis*. Acta Phytotax. Geobot. 38:33–44.

Thompson, J.D., D.G. Higgins, and T.J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22:4673–4680.